

A Malayan Element in some of the Languages of Southern Indo-China.

BY C. O. BLAGDEN.

In a former paper I endeavoured to point out that the aboriginal dialects of the Malay Peninsula show distinct traces of an Indo-Chinese element, impressed upon them, probably at a fairly early date, by the intrusion from Southern Indo-China of a race of Mon-Annam stock speaking a language which was closely allied to that of the Peguans and Cambojans.* The object of the present paper is to introduce the readers of this Journal to what may perhaps be appropriately described as the converse phenomenon, namely, the persistence (from a still remoter era) in some parts of Southern Indo-China, of distinct relics of an independent group of Malayan dialects, underlying the now dominant Indo-Chinese languages of that region.

As might be expected, the modern representatives of this group are far from being pure Malayan tongues: they exhibit obvious traces of the Mon-Annam and other influences to which they have for many centuries been subjected, and it is by no means certain that, in their present mixed condition, they can all claim to be classified in the Malayo-Polynesian family of languages. But whether that claim, which is sometimes made for them by French scholars more familiar with the Indo-Chinese than the Malayan languages, could be substantiated or not; whether, that is to say, these mixed dialects are to be regarded

* This subject has been learnedly and (so far as the materials at his disposal permitted) exhaustively handled by the Rev. Father W. Schmidt in a recent paper "Die Sprachen der Sakei und Semang auf Malacca und ihr Verhältnis zu den Mon-Khmet-Sprachen", which appeared in *Bijdragen tot de Taal-hand-en Volken-Kunde van Nederlandsch-Indië* Vol. LII (Series 6, part 8) Fasc. 3-4 (The Hague, 1901).

It remains to be seen whether the author's conclusions will stand the test of the further evidence that can be adduced; but at any rate he has marshalled the evidence that was before him with admirable skill and scientific acumen.

as genuine Malayan languages overlaid with foreign accretions, or, on the other hand, as alien tongues containing a large number of old Malayan loan words, is not for the present purpose very material. In order to decide this point and to determine whether these mixed languages partake more of the Malayan or of the Mon-Annam type, a careful study of their structure and grammar would be required, but the materials for such a study are at present very deficient, and in either case these dialects even in their present state presuppose, as I intend to show, the existence of a distinct Malayan continental group established at a very remote period in the south of Indo-China.

The chief of these languages is Cham, the language of the ancient Hindu kingdom of Champa, which in medieval times occupied the country now called Annam, and in the period just preceding its fall (which occurred in A. D. 1471) had its centre on the East coast of Indo-China about lat. 14° N., though one of its earlier capitals was as far north as lat. $17^{\circ} 37'$ N. This language is still spoken in a few inland villages of the Annamese province of Binh Thuan, near lat. 12° N., and by the emigrant Cham community in Camboja; the latter is now Muhammadan in its entirety, but the Chams that remain in Annam are mostly pagans. Each group has its own dialect, but apart from slight variations the language of both is the same. It is written in a complex alphabet of Indian origin: inscriptions, both in Sanskrit and in Cham, abound in Annam, and the former go back to about the 3rd century after our era.* According

* The Sanskrit inscriptions were dealt with in a paper "L'Ancien Royaume de Campa d'après les inscriptions" by M. Abel Bergaigne in the *Journal Asiatique* (Paris) Jan. Feb. 1888.

The inscriptions in Cham, which have more interest for us, from the Malayan point of view, than the Sanskrit ones, have been dealt with by M. Etienne Aymonier in a paper "Première Étude sur les Inscriptions Tchames," in the same journal, Jan. Feb. 1891. The earliest known of these Cham inscriptions dates from about the beginning of the 9th century A. D.

In an inscription dated a little later, recording the dedication of two fields to pious uses, the expression used is *huma dua nan*, lit. "fields two those"; the word for God is *Yang*, the old word which survives in Malay *kayangan* and *sembahyang*. Most of the rest of the inscription is full of Sanskrit words, as indeed the whole series

to Ptolemy the metropolis of this region was Balonga. This place can be clearly identified,* on other grounds besides mere similarity of name, with Bal-Angoué, of which the ruins situated near the coast about lat. 14° N are still in existence, and which was therefore apparently the first, or at least the earliest known, as it ultimately became the last, of the Cham capitals. Its fall is narrated, curiously enough, in the *Sějarah Malayu*, where it is called Bal, the generic Cham word for "metropolis" or "capital."

The Chams, in fact, are the remnants of what was once a highly civilized nation: they were the furthest outpost of Indian civilization on the Asiatic continent, and their country was a borderland where for over a thousand years Indian culture struggled with and was eventually vanquished by Chinese, the latter being represented by the Annamese, who though non-Chinese in origin had become civilized under Chinese tutelage.

Such is the history of the Chams in outline: but legends carry it back even further, for the Cambojan traditions, for what they are worth, represent the Chams as having been in occupation of Camboja when the Cambojans first arrived there, some centuries before the Christian era: the immigrant Cambojans are said to have intermingled at first with the Chams but eventually to have got the upper hand and driven out their king.

Physically the Chams appear to resemble the Malay and Indo-Chinese types, being described as somewhat fairer than the former. Some of them appear to show traces of Indian and Arab blood. Their language, of which a good grammar has been published, is in its present condition a mixed language containing a relatively large number of Mon-Annam elements. Some have regarded it as a Mon-Annam language saturated with Malayan loan words, others maintain that it is a Malayan language modified by Mon-Annam influences. As will appear in the sequel, I am not sure that this may not be something

of Cham inscriptions appear to be, the language in which they were written bearing much the same relation to the spoken Cham, as Kawi probably did to the contemporary spoken Javanese.

The series extends into the 15th century, to a few years before the fall of the kingdom.

* See J. R. A. S. (1899) 665.

like a distinction without a difference; but certain it is, at any rate, that Cham contains a very large percentage (perhaps nearly 50 per cent.) of pure Malayan words; and in this respect it seems to exceed its neighbours, the dialects to be next mentioned.

It is in the hilly country bounding Annam on the west and separating it from the valley of the Mekong River, about lat. 13° and 14° N., that these three dialects are found: they are spoken by three savage tribes called respectively Cancho, Rodê and Chréai. These tribes appear to be on much the same plane of civilization as the Orang Hutan of the South of the Malay Peninsula; their dialects are unwritten, and we owe such slight knowledge of them as we possess to the investigations of the three or four French explorers and administrators who have interested themselves in them. Practically that merely amounts to vocabularies of about 120 or 150 words of each of these dialects.* Besides these, there are other dialects in this region which are apparently more or less related to the above, and of some of which even less is known: † most of them however show decidedly more relationship with the Mon-Annam than with the Malayan family, the elements which they have in common with the latter decreasing in relative importance as one proceeds north and west from the old Cham region.

The only other dialect I propose to deal with here belongs to a different quarter altogether: it is spoken by the Selung (or Silung or Salone, as they are variously called) a sea-faring race who inhabit the numerous islands that fringe the Western Shore of Tenasserim (Lower Burma) from about lat. 13° N. to about lat. 10° N., and are marked on maps with the rather high-sounding title of the Mergui Archipelago.

These people may fairly enough be styled a distant branch of the Orang Laut. Their physical type, to judge from photographs, is more or less that of a rude Malayan race, with (possibly) some admixture of other elements, (of which the Indonesian may be one, as the Selungs, or at least some of them, are

* These are given in Moura, "Le Royaume du Cambodge."

† Of the Bahnar, however, a good dictionary by Dourisboure has been published (Hong Kong, 1889). It is a Mon-Annam dialect, but contains a certain number of Malayan words.

mesaticephalic, while the true Malays tend to the *brachycephalic* type). The three wild tribes previously mentioned, I should have said, appear from descriptions and such illustrations as I have seen, to be at least in part of non-Malayan stock: some authorities have insisted much upon their Caucasian type, by which I suppose is meant that they differ considerably from the Mongoloid type of features common to both Indo-Chines and Malays.

The Selungs, whatever their race may be, are pagans in a low state of civilization, and their language is an unwritten tongue. It comprises several dialects differing considerably from one another, so that people from two islands barely eighty miles apart have some difficulty in carrying on an intelligible conversation together. Several short vocabularies* of this language have been collected at various times by different persons, and they serve to illustrate these dialectic variations: but as it is not quite clear to which dialects they respectively refer, the Selung must for our purposes be dealt with as one language. It would appear to be really a Malayan language, less mixed with other elements than are the tongues already mentioned, and its claim to be mentioned here at all rests merely on its present geographical position: but being the speech of a sea-roving race of islanders it is obvious that its position does not furnish such cogent evidence for the antiquity of Malayan elements in Indo-China as do the inland dialects previously enumerated; nor is it as closely connected with any of them as they evidently are with one another.

It may however be said to form a link in the chain between these mainland dialects and languages of the Eastern Archipelago; and that is the reason why mention is made of it here, although its existence does not really affect the main argument of this paper.

It would be merely wearisome to present a whole series of vocabularies of the five languages I have enumerated: a few words will serve to convey some idea of the nature of the Malayan elements which they contain and will exhibit the

* They are given in Anderson, "The Selungs of the Mergui Archipelago."

peculiar character of their relation to the Malayo-Polynesian family of languages quite sufficiently for the present purpose.

The numerals, which are very characteristic, are as follows:—

	<i>Cham.</i>	<i>Cancho.</i>	<i>Rodê.</i>	<i>Chrêai.</i>	<i>Selung.</i>
One	thaa, sa	sa	sa	sa	chă, chet
Two	dvaa, dva	doa	doa	toa	twā
Three	klău	clou	to	clou	tahlow
Four	pak	pac	pac	pac	păt
Five	limœû	lema	ema	léma	lemah
Six	nam	nam	nam	nam	nam
Seven	tijuh	tuchu	cachu	tuchu	loojo
Eight	dalapan	salapan	sapan	repan	wahlow
Nine	{thalapan, salapan, samilan }	doalapan	doapan	toapan	chowai
Ten	{tha pluh, sa pluh }	saplu	plu	plu	taplaw
Eleven	saplul sa	saplu sa	plu sa	plu sa	{taplaw-chă taplaw-chet
Twelve	saplu dva	saplu doa	plu doa	plu toa	ta plaw-twa
Twenty	dva pluh	doa plu	doa plu	toa plu	twa plaw
Hundred	ratuh	[Not given]		retus	allataw
Thousand	ribău	[Not given]		ha repou	[appăn]

The *th*-* forms in Cham belong to the Binh Thuan, the *s*-forms to the Camboja, dialect. Presumably the double forms in Selung are also dialectic variants. The spelling of Selung is the old fashioned English, that of Cham the modern scientific system†; as to the rest, they are collected by French authorities but I am not quite clear on what system they are spelt.

These words are interesting as exhibiting a numeral system which, though unquestionably and obviously Malayan, is in some

* This *th*- is the English sound in *thing*. Some dialects of Achinese also turn *s*- into *th*- in this way.

† Slightly modified by the French tendencies of the transliterator. His *e* = *u* his *œ* = a sound varying between the vowels of Fr. *cœur* and *œu*, or the two *eu* in Fr. *leureux*. But *u* is the real *u* (Fr. *ou*); *œû* is a lengthening of *œ*.

respects clearly more archaic than that of Malay and could not, therefore, have been derived from it. In fact, even if these words were all that we knew of the dialects in question, we should be justified in saying that they constituted a distinct subgroup of languages, not directly derived from any existing Malayan group. The forms for *one*, *two*, *four*, *five* and *six* run practically through the whole Malayo-Polynesian family almost unchanged. In *four* the mainland dialects approximate most closely, perhaps, to the Bugis *äpak* and Madurese *ëmpak*, unless indeed the *-k*, which appears to be unpronounced in these two languages, is to be regarded merely as a device of writing, not as the remnant of a real *-k*; Selung agrees with the Javanese and Dayak *pat*. In *six* they all agree with the Javanese *nëm* in the absence of the first syllable of the word (Malay *anam*) but retain the *a* of the second syllable like the Malay (also the Madurese *änam*); the Achinese and Kayan Dayak form *nam* is identical.

The forms for *three* agree substantially amongst themselves and (except that some have a guttural for the initial *t*-) with the great majority of the Malayo-Polynesian family which retains the old form *tolu* or *tëlu*; but differ from Malay, which has another word, *tiga*. The nearest approximation to the Cham Cancho and Chréai forms appears to be the Bisaya (Philippines), *tló*: compare also the Sulu * *Kätluân* (= *Ka-tlu-an*), "thirty." For the guttural, compare Sulu *iklog*, Selung *k'loen*, with Tagalog *itlóg*, Malay *tëlor*, "egg." The Rodê contraction *to* recurs in Sulu.

The forms for *seven*, on the other hand, differ from the typical Malayo-Polynesian *pitu* and agree substantially with the Malay *tujoh*, save only that Selung puts *l*- for *t*-.

In both these cases, it is very noticeable that the dialects now under consideration agree substantially with Achinese (*tëlu* or *lu* pronounced *tëlhée* and *thée*, "three;" and *tujuh*, "seven") and with some of the Dayak dialects of Borneo, for which the reader may refer to No. 5 of this Journal, where out of a list of eleven dialects, ten have forms of *tolu* for *three*, and eight of those ten agree with some others not included in the ten in having forms of *tujoh* for *seven*.

* Between Borneo and the Philippines.

In *eight* and *nine* there is some confusion, which may be due either to the collector or to the wild tribes themselves; possibly the latter get a little mixed when they come to the higher numbers. Anyhow, they are said to use for *eight* a form *salapan* which occurs again in Sundanese (Java) and also in Mangkasar (Macassar, of Celebes), in the latter under the form *salapang*, and there means, as it ought to mean, *nine*. Oddly enough, the Minangkabau Malays use it, interchangeably with *dulapan* (*dĕlapan*), and also make it mean *eight*. Vice versa, these wild tribes use variants of the Malay and Achinese form of *eight* for *nine*. Cham, it is to be observed, uses both forms correctly, but has also another form for *nine*, viz., *Samilan*, the Malay *Sambilan* (*Sĕmbilon*), which may perhaps be merely a loan word from Malay itself.

There has been, in historical times, a Malay immigration from Sumatra (and particularly, it seems, from Minangkabau) into Camboja (where this form *Samilan* is used) and the Cham and Malay communities in that country, though distinct, are in close contact with each other, and being of one religion sometimes intermarry.

It is noticeable that Selung differs from the other dialects in having preserved, though in rather uncouth shape, the original Malayo-Polynesian forms for *eight* (*walu*) and *nine* (*siwa*).

In the forms for *ten* these dialects agree substantially with the Achinese *pĕluh*, in shortening the first syllable; this does not, apparently, occur in the Bornean dialects, which in other respects show a fairly close resemblance in their numeral systems.

For *eleven* and upwards the dialects agree amongst themselves and with some of the Bornean dialects, but differ from Malay, Achinese, Javanese, etc., in not using forms compounded with *-bĕlas* (originally *-walas*, the Malay *balas*, "to repay," with the meaning "to return," i.e. to the hand on which the counting was first began).

The Selung for "hundred" apparently has the prefix *sa-* "one" reduced to *a*, which occurs also in a Cham subdialect as *ha-*. For the *-l-* of Selung *yahloam*, Malay *jarum*, "needle."

Thus while there are here particular words agreeing, each with some different Malayan language or group of languages,

the sum total of the numeral system of these dialects is quite characteristic in its individuality.

A similar state of things prevails in regard to many other common words, as the following specimens will suffice to show:—

	Cham.	Cancho.	Rodè.	Chrèai.	Selung.
Dog :	<i>athân.</i>	<i>ason.</i>	<i>so.</i>	<i>so.</i>	<i>oïee, aai.</i>

Melano-Dayak *asan* comes nearest but the word, though not found in Malay (except in the expression *gigi asu*, "canine teeth") is very wide spread, e.g. Javanese *asu*.

Fowl :	<i>menuk.</i>	<i>menuc.</i>	<i>menuc.</i>	[<i>tus</i>].	{ <i>manok.</i> <i>mayauak.</i>
--------	---------------	---------------	---------------	-----------------	------------------------------------

Compare the Javanese (and almost universal Malayo-Polynesian) *manuk*.

Tiger : *rimong.* *remong.* *imong.* *lemong.* ———

(The Selung word is different, viz : *paumoo, puuk*, which finds its analogues in aboriginal dialects of the Malay Peninsula, e.g., Tëmbe' *ma'nu* (for which see No. 24 of this Journal, p. 17). The Achinese form is *rimong* like the Cham. I think there is no reason to doubt the identity of the word with the Malay *rimau*. Possibly the form *hariman* is a sort of Hobson-Jobson word, that is to say, really the old native Malayan word for "tiger" but twisted into its present form by a fanciful notion that it ought to mean "the beast of Hari" (*harimriga*, see Maxwell, Manual of Malay, p. 21). I confess that even Sir William's brilliant scholarship cannot convince me that his Tamil "male lion" derivation is the right one.

Elephant : *liman.* *eman.* *romon.* *lomon.* ———

(Selung has *gazah*, the Malay *gajah*, a word of Sanskrit origin). Compare the Bulud Opie (Borneo), Javanese and Lampong (Sumatra) *liman* : this word, which is not found in Malay or Achinese, is probably derived from *lima*, the old word for "hand," the application being to the end of the animal's trunk. One of the Sanskrit names for the elephant (*hastin*) has a similar derivation ; and compare also his Latin epithet *anguimanus*, "having a serpent for a hand."

Ox, cow: *tamov. lemo. imo. romo. l'mn*: Malay *lěmbu*, in Achinese the same, and also *lěmo*.

Rain: *hujan. ujan. hayan. yan.* $\left\{ \begin{array}{l} kujan. \\ kuian. \end{array} \right.$

Malay *hujan*: but Batak and Javanese *údan**, Tagalog and Bisaya *olan*. Selung *k-* represents Malay *h-* in *ketam* (= *hitam*), "black" and a few other words.

Root: in Cham *ugha, agha* (in accordance with the peculiarity referred to below): this is not, apparently the Malay *akar* but *urat*, "Sinew." In form it is nearer to the Formosan *ugat*; Tagalog and Bisaya *ogát* than to any other forms. Batak in this word agrees with Malay.

In a sub-dialect of Cham of which specimens are given by Morice in an article entitled "Les Tiams et les Stiengs" in the "Revue de Linguistique" Vol. VII, vii, pp. 359-370, *r-* is often re-placed by *g-* e.g. *agopao* (= *saribu*) a "thousand": *hagaton* (= *saratus*), "a hundred." In Tagalog these words appear as *libo* and *gatós* respectively.

Tongue: in Cham *dilah, dalah* (both being used); approaching nearer to the Tagalog *dita*, Bisaya *dila*, than to the Malay and Achinese *lidah*. Batak also has *dila*: here, therefore, Cham agrees closely with Batak and the Philippine languages but differs from Malay and Achinese.

Belly: $\left\{ \begin{array}{l} téan. \\ tjan. \end{array} \right.$ *téan. téan. kajéan. k'lan.*

Bisaya, Iranun and Dusun *tian*, Sulu *tián*. *Tian* is given in some Malay dictionaries as a Javanese loan word meaning "belly (of a pregnant woman)." In Achinese *tijěn* means "fœtus," *mětijěn* "to be pregnant"; in Cham *mætéan* means "pregnancy," *boh téan* (literally "fruit of the belly," Malay *buah*, Javanese *woh*, "fruit") means "family."

Hand: *tangin. tengam. cangan. tangin. lengan.*
Malay *tangan*, Dusun *lángan*, Dusun of Kimanis *longon*.
For the Selung *l-* = Malay *t-*, compare *loojoo* (= *tujoh*), "seven."

* Van der Tuuk's third rule: "when a *j* of Balinese and Malay is *d* in Batak, the Javanese and Kawi both also have *d*."

Cham.	Cancho.	Rodê.	Chérai.	Selung.
Nose : <i>adung.</i>	[<i>chnu</i>].	<i>dung.</i>	<i>dung.</i>	<i>{yoong.</i> <i>{yong.</i>

Malay and Achinese have *hidung*. Cham uses both *adung* and *idung*. Compare the Tidung (Borneo) *adung*, Dusun of Kimanis *adung*. Javanese and most of the Bornean dialects replace this *d* by *r*; the Philippine languages (and in this word Madurese also) have *-l-* here; Batak has *-g-*. The importance of this particular set of consonantal correspondences was also first pointed out by the late Dr. H. N. van der Tuuk. They constitute his second rule:—"When the Malay and Balinese *d* of equivalent words is represented by *l* in Bisaya or Tagal, both the Javanese and Kawi have *r*." *Chnu* is probably Cambojan.

Fire : <i>aprēi.</i>	}	<i>apui.</i>	<i>pui.</i>	<i>puoi.</i>	{ <i>appoi.</i> { <i>apoi.</i> { <i>apoe.</i>
<i>apuei.</i>					
<i>apui.</i>					
Malay <i>api</i> , but Achinese and several Dayak dialects, etc., have <i>apui</i> .					

Water : <i>iā,</i>	}	<i>ea.</i>	<i>ea.</i>	<i>ja.</i>	<i>{awên.</i> <i>{awaen.</i>
<i>ear.</i>					

Malay *ayêr*, Achinese *iyêr*, Madurese *aeng*, etc.

Stone : <i>batāu.</i>	<i>pétou.</i>	<i>bato.</i>	<i>potou.</i>	<i>batoe.</i>
-----------------------	---------------	--------------	---------------	---------------

Malay *batu*, the Achinese equivalent is written in the same way but pronounced *batée*.

The few words here given suffice to show that these dialects have peculiar points of relationship with several widely separated Malayan groups of languages and could not have been derived from any one of them. Their affinities appear to be most marked with Achinese, as is shown especially by the fact that in common with that language (and quite the opposite to Malay), they tend to throw the accent on the last syllable, which is consequently often strengthened to a diphthong, at the expense of the first, which is weakened and sometimes entirely suppressed: Compare *pluh*, "ten" with the Achinese *pěluh* and contrast the Malay *puloh*; similarly compare the forms, in

Cham, Achinese and Malay respectively, **thun*, *těhun*, *tahun*, "year"; **dhan*, *děhěn*, *dahan*, "bough"; *ngan*, *ngon*, *děngan*, "with"; *doł*, *duk*, *dulok*, "remain, dwell, sit"; and *mətai*, *maté*, *mati*, "dead." Selung has *matai*, which form also occurs in Bornean dialects as *materi*.

It is probably owing to the same tendency to weaken the first syllable, that Cham has *hajan* for *hujan*, "rain," *akan* for *ikan*, "fish," *adung* for *hidung*, "nose," *balau* for *bulu*, "hair," and the like: and here it goes further in this direction than Achinese or any other Malayan language that I am aware of, although this vowel change appears also (but more rarely in some Bornean dialects, e.g. Tidung *adung*, "nose," Biadju Dayak *balau*, Lawangan *balu*, Siang *warlo* † [sic], "hair.")

It will of course be understood that the words here given have been expressly chosen with a view to exhibiting the Malayan element in these dialects, and that alien, especially Mon-Annam forms have been deliberately avoided. The Malayan element is strongest in the substantives, but is also represented in some of the verbs and adjectives, e.g.

	Cham.	Cancho.	Rodê.	Chréai.
Buy :	<i>blěi</i> .	<i>bloi</i> .	<i>bloi</i> .	<i>bloi</i> .
	Malay <i>běli</i> ,	Achinese, <i>bloi</i> .		
Sell :	<i>pablěi</i> (in Cham :	the rest are different) :	Achinese <i>publoi</i> .	
Give :	<i>brěi</i> .	<i>brey</i> .	<i>broi</i> .	<i>proi</i> .
	Malay <i>běri</i> ,	Achinese <i>bri</i> .		
Descend :	<i>trun</i> .	<i>trunh</i> .	<i>trun</i> .	[<i>tumau</i> .]
	Malay <i>turun</i> ,	Achinese <i>trun</i> .		
White :	<i>patih</i> (Cham) :	<i>potajak</i> , <i>patuik</i> (Seiung) :	Malay <i>puteh</i> .	
Drunk :	<i>məbuk</i> (Cham) :	Malay <i>mabok</i> .		
New :	<i>barūv</i> (Cham) :	Malay <i>bāharu</i> .		
Unripe :	<i>mətah</i> (Cham) :	Malay <i>měntah</i> .‡		

* This is a different *th*- from the other: this *th*- and *dh*- are true aspirates.

† I take these examples from C. den Hamer's *Proeve van eener Verglikende Woordenlijst van zes in de Z. O. Afd. v. Borneo voorkomende Taaltakken*.

‡ For the present purpose it is not necessary to pursue this comparison further. Suffice it to say that the Malayan element can be traced (at least in Cham and to some extent in Selung, there being no

The main object of this paper being merely to point out the existence of Malayo-Polynesian words in these languages and not to determine the difficult question of their right to be classified as genuine members of that family, I shall pass somewhat lightly over their grammatical characteristics of which indeed, except as regards Cham, little is as yet known. Cham forms its derivative words, like the Malayan, but unfortunately, also like the Southern Mon-Annam languages, with prefixes and infixes: The common ones in Cham are the prefixes: *pa*, *mæ*-, *ta*- or *da*- and infixes: *-an*-, *-næ*- and *-am*- or *-mæ*-. Most of these reappear, in more or less similar forms, with much the same force, in Achinese; but also in Cambodian, where they are very freely used, and to some extent in Peguan.* Suffixes, corresponding to the Malay *-kan* and *-an*

	<i>Prefixes.</i>	<i>Achinese.</i>	<i>Cham.</i>	<i>Khmer.</i>	<i>Mon.</i>
Verbs of action: causal or merely transitive	...	pê-, pu-	pa-	p-, ph-	p-, ph-, b-
Verbs, generally intransitive	...	mě-, mu-	mæ-	?	mă-
<i>Infixes.</i>					
Verbs of state, intransitive	...	-êm-	-mæ-	?	-m-
Substantives	...	?	-mæ-	-m-, -amn-	-m-
Substantives	...	-ên-	-an-	-n-, -an-	[-an- ?]

In some other cases, where the forms agree, the meanings appear to differ somewhat.

do not appear to be in use at the present time either in Cham, Achinese, Cambodian or Peguan; but if the derivation given above for *limæn* (*liman*) from *lima* is right, they must have existed formerly to some extent in Cham.

The Selung dialect forms verbs by prefixing *me*- as in *metoqam*, "to smell" (Malay *cium*), *na*- as in *na-baut*, "to make" (Malay *buat*), *naleat*, "to look" (Malay *lihat*), *nadök*, "to sit" (Malay *duduk*, Achinese *duk*, Cham *dok*); also, apparently, by nasalizing the initial consonant, as in *nadone*, "to sleep" (Malay *tidor*) and *nakoat* "to fear" (Malay *takut*). But

data for the other dialects) through most of the parts of speech, but the non-Malayan element is also, apparently, present in them.

* A few instances of this general correspondence must suffice: there are of course many differences in detail.

this last may possibly be due to the phonetic decay of a prefix of the form *man-* or *měn-* (the Malay *mě-*, *měng-*, etc.): for a word like *mangai*, "to cry" seems to presuppose an earlier *manangai* (Malay *tangis*, *měnanngis*) and *mawah*, "to laugh" an earlier *manawah* (Malay *těr-tawa*). The loss of a medial *-n-* seems more probable than that of a *-t-*: it may be, however, that the Selung in these words as in "seven" had replaced the *t* by *l*. In that case these forms probably exemplify the prefix *me-* above.

Selung has the suffix *-kan* e.g. in the word *makkān* (for *ma-bahkan*, *ām-bahkan* or *měmbahkan*, from *bah*, to "bring," Malay *bawa*).

The ideological order of these languages is unknown to me, except that in Cham (as in the Mon-Annam languages again) it appears to agree substantially with the Malay order: the attributive adjective and the genitive follow the principal noun, the object follows and the subject precedes the verb; but in Selung the object precedes the verb, which is very strange, unless it is due to the sentences having been collected through the medium of a Burmese interpreter, in speaking to whom the Selungs may have cast their words into the Burmese order. It is curious that Andamanese exhibits the same phenomenon: but there is no evidence that the Selungs are in any way connected with the Andaman islanders: both in physique and in language the two races are quite distinct from one another.

I have already indicated the conclusion to which a necessarily rather superficial comparison of these dialects seems to me to point; I regard them, or at least all of them except Selung, as proof positive of the establishment on the mainland of Southern Indo-China of a Malayan sub-family which must date its separate existence from a period so remote as to be coeval with the differentiation and dispersal of the existing insular language groups of at least the Western part of the Malayan Archipelago, and which formed something like a link between the Sumatra, Bornean and Philippine groups.

I think it is worth adding that the southern Mon-Annam languages, which so closely resemble the Malayan in certain of their structural forms, though by far the greater part of their vocabulary is radically different and non-Malayan, owe this

resemblance, in my opinion, to the fact of their having developed on what I believe was originally a Malayan soil. The true explanation of the peculiarities which they share in common with the Malayo-Polynesian family is, I believe, that they have been formed by the synthesis of a language introduced by alien immigrants from the north with the Malayan speech of a people who then already occupied Southern Indo-China. The northern invaders must have absorbed and assimilated these primitive Malayo-Polynesians and imposed upon them their alien language, which in its turn has been twisted, in the mouths of their mixed descendants, into something of a Malayo-Polynesian form, by a process that has been aptly called "inverse attraction."

The result of such an introduction of a strange tongue is, as a rule, that it becomes modified or recast into some form that comes natural to the people upon whom it is imposed: this may be illustrated by such well known cases as the Pidgin English, of the China ports, Negro English, or the Malay of many Chinese, Tamils and Europeans.

In such cases the mere vocabulary, though foreign to the speaker, is learnt readily enough; but he cannot help speaking his new tongue in the manner of his old one. He pronounces the new words in the way that comes easiest to him and utters them in what is to him the natural order, though that may not be the order proper to the language as spoken by those whose original speech it was. If it was natural to him to use prefixes and infixes in his old language, I imagine he would be apt to apply them to his acquired tongue in the same way and for the same purposes. This, to my mind, is the explanation of the curious fact that in Cambojan and Peguan we find these modes of formation, which are so characteristic of the Malayo-Polynesian family, while the difference of the material elements of language, i.e. the words themselves, prevents us from admitting an original kinship between the Mon-Annam and the Malayan families of speech.

I am afraid that this idea of the formal elements of language surviving, while the native vocabulary is gradually being superseded by foreign words, may remind some people of the persistence of the grin after the disappearance of the Cheshire

cat. But the real analogy is to be found in those petrifications where every cell and fibre of the original wood or other substance are in course of time accurately reproduced by the stony deposit that replaces them. To drop figures of speech, which, however apt, can never be conclusive, when one considers that the Malayan languages readily adopt foreign words and instinctively fit them up with Malayan prefixes and suffixes, one can almost see the beginnings of such a process as I have indicated: words like *ka-raja-an*, *běr-akul* or even *di-rěport-kan* (which last can be heard any day when a Malay police officer reads from his Station report book in a Police Court) are instances taken at random, where a Sanskrit, Arabic or English loan word has been subjected to this treatment.

One has only to carry the idea out to its logical conclusion and imagine a Malayan language gradually allowing its native vocabulary to be superseded, more or less completely, by foreign loan words, and the result would be much the same as what we now find in southern Indo-China. If the process were arrested half-way, a fairly evenly mixed vocabulary would be formed, like that of Cham; a more advanced stage of change would result in something like Cambojan; while a thorough application of the same principle might end in producing a language like Peguan, where only a very small percentage of words is to be found which show any signs of kinship with the Malayan family. Nevertheless the ideological order of these languages, that is to say the order of words in a sentence, is substantially the same as in the Malayan languages* and the same system of prefixes and infixes (though not, apparently, of suffixes) still survives.

On the other hand a strong tendency is noticeable, of which it has been shown that even Achinese (Malayan language) exhibited the beginnings, to contract disyllabic words into monosyllables or at least into quasi-monosyllables, in which one of the two syllables is almost suppressed. There are also other

* There is reason to believe that in this respect the Mon-Annam languages did not differ originally from the Malayan.

peculiarities which distinguish the Mon-Annam from the Malayan group, e.g., a preference for hard sounds * (surds) and the occurrence of true aspirated consonants: these latter characteristics may be due to the non-Malayan element in these languages.

The hypothesis here put forward would account for the remarkable resemblance in structure and formal elements between the Malayan and the Mon-Annam languages, a resemblance which, so far as I know, no one has yet satisfactorily explained.† But of course it must remain a mere hypothesis until these languages have been thoroughly studied and compared with one another.

This much, however, is certain: one Mon-Annam language which cannot be accused of having been developed on Malayan soil, namely the Annamese, which grew up on the borders of Kwang Si, within the Chinese sphere of influence, does not exhibit these phenomena, but follows the Chinese system of tones, though it has not adopted the Chinese ideological order. I take it that the differences between Peguan and Cambojan on the one side and Annamese on the other are the measure of the difference between a Chinese and a Malayan environment.

Whether, however, this suggested explanation be the true one or not, there remains the fact that in Peguan, and still more in Cambojan, there are a fair number of words (too many to be due to accidental coincidence) which correspond in form with Malayan words of similar meanings. As already stated, they are generally more or less contracted or mutilated, by the weakening or entire loss of one syllable, while the Malayan languages retain them in their fuller disyllabic forms. That being the case, the presumption is that they are genuine Malayan words; and this presumption is strengthened when any of

* Clearly, however, it is at a relatively modern date that the Mon-Annam languages have changed some of their sonants into surds: for in many cases (especially in many of the Indian and some of the Malayan loan-words) they still appear as sonants in the written language. Conversely Cambojan pronounces some surds as sonants.

† Mr. Himly in his paper referred to below, throws out a hint that some such explanation is possible, but does not enlarge upon it.

them are found to occur again in some distant island dialect of the Malayan family.

I propose to give a few instances to show the forms which such words assume in Cambojan and Peguan, but before doing so, I may as well point out that Indian loan-words, as to the origin and derivation of which there can be no doubt, undergo a similar mutilation in the Southern Indo-Chinese languages so that an analysis of the changes exhibited by these Indian words will serve as a guide in identifying the Malayan words to be found in those languages, which are often hardly recognizable without some such help.

The following are examples of words of Indian origin common to Malay and these two languages: I give the Malay, rather than the Sanskrit form, because the former is more familiar to those who know Malay.

<i>Malay.</i>	<i>Cambojan.</i>	<i>Peguan.</i>
Kala	... kāl kāla.
Kēchapi	... chāpey [chapěy]	... —
Guru	... grūw [Krū] —
Chandra	... chand [chān] —
Jambu	... jāmbūw [chōmpū	... —
Dewata	{ dew-ta [tévoda] deb-ta [tépoda]	{ dewatan [tewātau].
Dosa	... dōs [tōus] duh [tuh].
Nagara	... nagar [nokor]	... —
Naga	... nāg [néak] nāk [naik].
Puasa	... puos [buos] —
Bangsa	{ wangs [vong] pangs [pong] ...	{ wang [weang] wongsa.
Muka	... mukkh [mūkh]	... muk.
Raja	... rāj [réach] rājā [reachea].
Satwa	... satw [sāt] sat [sāt].
Sutra	... sūt [saut] sut.

The following list shews some of the similar changes which Malayan words suffer, viz.

I. Suppression or weakening of the first syllable:—

Malay.		Cambojan.		Peguan.
Kayu	jhéc [chœu]	... chhu [tsu]
Kijang	k-tân [kědan]	... ———
Katup	k tãp [kědap]	... ———
Garam	krām [kram]	... ———
Jawa	jwā [chvéa]	... ———
Tarum	trâm [trôm]	... ———
Pusat	phchêt	... ———
Perak*	prāk [prāk]	... ———
Běsi*	————	Pāsoa.
Sarong	srôm	... ———

II. Loss of initial consonant:—

Chin chin	ânchién	... kâchin.
Těbu	ainbau [âmpou]	... bau.
Tabong	ainbang [ampong]	... ———
Daching	anjing [ânching]	... ———

III. Loss of first syllable:—

Tumbok	pok [bok]	... ———
Abang	pōng [bōng]	.. ———
Lětak	tāk [dăk]	... ———
Kěring	rīng	... ———
Esok	sâék	... ———

*[Note] Achinese *běsoi*, “iron.” It is perhaps worth noticing that the Cambojan word for *silver*, like the Peguan for *iron*, is Malayan, while the Cambojan for *iron*, viz., *těk* [děk] is common to it and Chinese. The Cambojan word for *gold* is *mas* [néas]; the same as the Malay *mas*, *āmas*; but this is believed to be of Indian origin. For *tin* the Peguans use the expression *pāsoa dāk* [pāsoa daik], literally “water iron,” alluding presumably to the alluvial formations where tin ore is got by washing river sand, while the Cambojans call it *Samna pāhāng* [Sămnă pahăng], from which, as *samna* appears properly to mean “lead,” I conjecture that the Cambojans first got their tin from Pahang, for the word *pāhāng* does not seem to have any meaning in their language, so far as can be ascertained from the Dictionary. Similarly in some of the Western languages (e.g. Arabic and also Hindustani, I believe) tin is called by a name *al-kalā'i* derived from Kālāh, a place on the Western shore of the Peninsula probably identical with Kēdah.

IV. Loss of second syllable :—

Pěchah	...	pek [bék]	..	pākaw [pāko].
Patah	...	Pāk [bak]	...	puīt [pat].
Buka	...	pěčk [bòk]	...	pāk.
Mata	...	——	...	mat [mât, mot].
Tanda	...	tān [dan]	...	——
Tolak	...	tol [dol]	...	——
Pakai	...	bāk [peak]	...	buik [puk].

The Cambojan and Peguan words have been transliterated, to the best of my ability, from the written languages: where the pronunciation is different, this is indicated by a second form in square brackets, following in the case of Cambojan, M. Aymonier's spelling and in the case of Peguan the indications given by Haswell, adapted to the ordinary modern system of romanization.

This list could be considerably lengthened, specially as regards Cambojan, if space permitted: but I think it is enough * to show that there is a field of research waiting for any Malay scholar who has a fancy for hunting up Malayan words in these languages. It would however be a great mistake to suppose that the bulk of the vocabulary of Peguan or Cambojan can be accounted for in this way: the contrary is the fact, and at first sight any Malay student looking through a dictionary of either of these tongues would be struck with their non-Malayan aspect. It is by neglecting the essential relationship which exists between Peguan and Cambojan † and ignoring the

*In presenting a list which merely compares a few words in Peguan and Cambojan with what I believe to be the corresponding words in Malay, without taking into account the other Mon-Annam dialects and the other languages of the Malayan family, I am aware that I am offending against one of the primary canons of comparative philology. But my present object being to make out merely such a *prima facie* case as will justify further investigation in this direction, I have thought it superfluous to bring in the corroborative evidence that can be supplied from the other languages. I hope some day to deal with this matter more fully and systematically.

† It will interest Straits readers to know that this was first noticed by our Straits authority, J. R. Logan. It has since been conclusively proved by Forbes in his "Languages of Further India."

wide differences in lexicographical material between the latter and the Malayan languages, that some authorities have been misled into denying the existence of a Mon-Annam family and asserting that Cambojan should be classified as a member of the Malayan group.

So far as it goes, this list of words serves to illustrate the subject of this paper by giving another instance of the traces of a Malayan influence in Indo-China, which must be of very ancient date, and which is obviously an important element to be considered in relation to the unsolved problem of the origin of the Malayan races.

Many considerations point to the conclusion that at least some part of the ancestry of those races* is of continental Asiatic origin: there are anthropological reasons, which I am unable to deal with, but which have been summed up roughly (and not very accurately) in the phrase "Mongoloid type;" ethnographical considerations, such as were dwelt upon by the late Sir Henry Yule † and others, specially a curious agreement between the races of the Archipelago and those of Indo-China in a considerable number of points of detail regarding customs and usages (a kind of evidence, which though very weak if depending merely on one or two points of agreement, is in its nature cumulative and gains strength in an increased ratio as additional points are discovered); and, finally, there is the linguistic evidence, the investigation of which is, however, involved in many preliminary difficulties. It is to be feared, for instance, that the late Mr. J. R. Logan's achievements in this direction are not a safe basis for further enquiries to start from. On the other hand Professor Kern,‡ by a comparison of

* I refer here more particularly to the true Malayan races inhabiting the western half of the Indian Archipelago, to whom alone the anthropological argument applies. How it is that the totally distinct stocks known as Papuan, Polynesian, Micronesian, etc., come to speak languages that cannot be severed from the Malayan family, is another problem, also at present awaiting solution. There seems, however, no doubt that it is the case, in spite of the difficulty of finding an explanation for it.

† Journal of the Anthropological Institute, 1880.

‡ In the paper to which a reference will be found below, the most conclusive, perhaps, of these words are the names for sugar-cane,

a considerable number of names of plants, animals and the like, which run (more or less) through the whole range of Malayo-Polynesian languages from Madagascar to Hawaii and from Formosa to New Zealand, has shown that the speakers (whoever they were) of the mother tongue from which these innumerable languages were evolved, were a seafaring people, of some moderate degree of civilization, (they were acquainted with the use of iron), who at the stage preceding the differentiation of these languages (but not necessarily originally) inhabited a long coastline of some good-sized country situated within the tropics, somewhere in the western half of the vast region over which these languages now extend. He points to the South-Eastern coast of Indo-China as the country that fits in best with this conclusion; and without going into details, lays some stress on the considerable Malayan element that is to be found in the existing languages of that region, which fact, as he observes, in view of the relative unimportance of the small Malayan communities to be found there in modern times, can only be explained by the hypothesis that they formerly constituted a much more numerous and powerful factor there than they do in our own day.

This last point it has been my endeavour to illustrate in the present paper.

It may be convenient if I summarize the conclusions to which the considerations here brought together appear to me to lead:—

(1) The Malayan element in Cham and its cognate dialects was not borrowed from any other Malayan language or group of languages. It has been separated from the western insular groups for as many centuries, as they have been from one another, and has become differentiated from them as they have amongst themselves.

(2) The Southern Mon-Annam languages and Cham are at once Malayan and non-Malayan: largely Malayan in structural formation, mixed but predominantly non-Malayan in vocabulary, they are probably the result of an intimate mixture between

banana, rice (in husk and husked), shark, prawn, sea-turtle, buffalo and crocodile: but there are a good many more besides.

Malayan and alien tongues. The Malayan element is strongest in the southeast, weakening progressively towards the north and west.

(3) At a remote age, before the introduction of the alien element just referred to, probably the whole coast of southern Indo-China from the Irrawady to the borders of Tongking, and certainly the eastern part of it from Cape St. James to the neighbourhood of Hué, was more or less occupied by communities speaking a pure Malayan language, possibly already slightly differentiated into dialects.

(4) It was probably from this region at a time when it was still purely Malayan, that the various emigrations took place, which ultimately carried dialects of that language to the distant islands in which they are now spoken.

I am content to rest this last proposition on the grounds put forward by Professor Kern in the essay already referred to; the other three appear to me to follow, though not all with the same degree of certainty, from the linguistic evidence of which some specimens have here been brought together.

Since writing the above, I have seen in the *T'oung Pao* for March, 1901 (Series II, Vol. 2, No. 1, p. 86) a review by M. Gustave Schlegel of a recent Siamese grammar. In noticing this work (which appears to be the best Siamese grammar hitherto published) after pointing out, what has been pointed out before, notably by the late M. Terrien de la Couperie, that Siamese contains a very large percentage* of words common to it and Chinese (especially, the numerals† which are, up to a certain point, pure Chinese loan words) and also a considerable number of Sanskrit and other Indian words, the eminent Chinese Professor of Leyden hazards the view that the residuum of Siamese will be found to be a Malayan language, and supports this thesis by a few words which no doubt are Malayan but may very well be loan words like the Indian ones; everything that the venerable professor writes is worthy of consideration, but

* De la Couperie puts it as high as 33½ per cent: "Languages of China before the Chinese" pp. 59-60.

† Not however, "one" and "two."

with all deference, I venture to say that this is indeed a bold theory. His chief argument, apparently, apart from the afore-said Malay loan words, is that Fu-nan (or Pu-nam), the old name for the country now called Siam, is capable of being explained by a Siamese derivation which M. Schlegel invents for it: unfortunately all monosyllabic languages lend themselves only too easily to hypothetical derivations of that kind; and that its people, in the early centuries of the Christian era, are described by Chinese chroniclers as being "ugly and black" with "curled hair," resembling, the Professor himself says, the Semangs. On the strength of this he assumes the Siamese to be Malayan. Everyone who has been to the Far East should know, and M. Schlegel can hardly have forgotten, that the Siamese are several shades fairer and the Semangs several shades darker than the average Malay complexion: and that neither Siamese nor Malays have curled or curly hair. His argument compels M. Schlegel to deny the historically certain fact that the Thai, that is the present Siamese, are comparatively recent arrivals from the interior of Northern Indo-China; and he entirely overlooks the essential unity of their language with that of the Laos, Shans, etc., right away to the Khamti on the eastern border of Assam and a string of tribes in southwestern China. If the Siamese spoken to-day at Bangkok is at bottom a Malayan language, so must be the languages of all these northern tribes, for they are substantially the same and cannot be severed from one another. That appears to me to be an exceedingly large conclusion to draw from a few Malay loan words to be found in modern Siamese, and I am convinced that it will be repudiated both by Siamese and Malay Scholars with tolerable unanimity.

Of course the possibility that there is a Malayan element in the blood of the modern Siamese of the South is not thereby excluded: that there should be such an element is an almost necessary consequence if the main argument of the foregoing paper has anything in it. But apart from modern intermixture which the difference of religions keeps at a minimum, it can only have come in at second hand through the Peguan or Cambojan inhabitants who occupied Siam before the Thai conquered it. That, however, is a very different matter from the

hazardous assertion that Siamese is a Malayan language, an assertion which requires far more cogent evidence to justify it than M. Schlegel has supplied in the article to which I refer.

It is hardly necessary for me to add that this paper is merely intended to draw the attention of the readers of this Journal to the subject; so far as the greater part of it is concerned, no claim is made for originality, and it is in the main merely a restatement of what has been set forth elsewhere in fuller form by others. My excuse for offering it to the Society is that some of the readers of this Journal may not have had access to the existing literature on the subject. At the risk of appearing egotistical, I desire to put on record that at the time my former paper was published, I had not heard of Professor Kuhn's admirable essay entitled "*Beiträge Zur Sprachen Kunde Hinterindiens.*" In it most of my conclusions were anticipated, and, if I had known of its existence, my paper would not have appeared, without at least some reference to it. The occasion for this personal explanation, which ought perhaps to have been made sooner, is a remark by Dr. Luering in No. 35 of this Journal.

I append a list of the principal authorities consulted:—

AYMONIER, *Grammaire de la langue Chame; Les Tchames et leurs Religions* (*Revue de l'Histoire des Religions*, 1891); *The History of Tchampa* (*Imperial and Asiatic Quarterly Review*, July 1893, and *Transactions of the Ninth Congress of Orientalists*, London, 1893); *Vocabulaire Cambodgien-Français; Dictionnaire Khmër-Français*;

VAN LANGEN, *Atjehsche Taal* (*Handleiding; Woordenboek*);

HASWELL, *Grammatical Notes and Vocabulary of the Peguan Language*;

STEVENS, *Vocabulary of English and Peguan*;

KERN, *De Fidjitaal; Taalkundige gegevens ter bepaling van het Stamland der Maleisch-Polynesische Volken* (*Verlagen en Mededeelingen der Koninkliken Akademie van Wetenschappen, Afd. Letterkunde*, Amsterdam, 1889).

MOURA, *Le Royaume du Cambodge*;

- ANDERSON, The Selungs of the Mergui Archipelago ;
- KUHN, Beiträge zur Sprachenkunde Hinterindiens (Sitzungsberichte der Philosophisch-, Philologisch- und Historischen classe der Königlichen Akademie der Wissenschaften ; Munich, 1889) ;
- HIMBY, Bemerkungen über die Wortbildung des Mon (*ibid.*) ; Ueber den Wörterschatz der Tscham-Sprache (*ibid.*), 1890.
- NIEMANN, Bijdrage tot de Kennis van der Verhouding van het Tjam tot de Talen van Indonesie (Bijdragen tot de Taal-, Landen Volkenkunde van Nederlandsch Indie, Leyden, 1891.)
- BRANDES, Bijdrage tot de Vergelijkende Klankleer der Westersche Afdeeling van de Maleisch-Polynesische Taalfamilie.
- VAN DER TUUK, Outlines of a Grammar of the Malagasy Language (Reprinted, from the Journal of the Royal Asiatic Society, in Vol. I of the Second Series of Essays Relating to Indo-China, for the Straits Branch).